

Identification and Classification Method of Agricultural Diseases and Insect Pests Based on Yolov3

Yuanbing LI¹, Bingkuan LI, Liang HE, Lulu FENG, Haotian LI

School of information engineering Sichuan agricultural university Sichuan, China

Abstract. This paper mainly selects four diseases and pests of corn for image recognition and classification, selects wavelet neural network algorithm for image processing, and then uses yolov3 neural network for image loss iteration to achieve better results. Through this research, we can better understand the application and depth of convolutional neural network in the field of image recognition. In this paper, the relevant data sets are used for neural network operation, and the data sets are processed appropriately. After the data set is obtained, the correlation degree is deleted and the wavelet algorithm is used for denoising to obtain the denoised picture and algorithm, which is equivalent to modifying and fitting the picture. Select the appropriate convolutional neural network on the existing basis, modify the convolutional neural network into several different neural networks, and then apply the convolutional neural network to the data set to obtain the classification effect after passing through the neural network and achieve the classification effect. At this time, the accuracy of the effect after continuous optimization can reach%, and achieve the corresponding effect, It is consistent with the expected effect.

Keywords. Convolutional neural network, image recognition, wavelet algorithm, prediction algorithm

1. Introduction

Due to the diversity and timeliness of crop diseases and pests, different kinds of diseases and pests are easy to occur under different conditions, resulting in a sharp decline in crop yield. Therefore, the prevention and elimination of crop diseases and pests has become an important link in agricultural production activities. We are faced with the rapid growth of population and the rapid reduction of cultivated land. Today's cultivated land is mostly occupied by industrial land and other land, which is of great significance to ensure food production and maintain the stability and controllability of plant growth cycle. In recent years, with the continuous development of technology, more than 500 kinds of pests have been found in China, including 30 common species in northern areas, such as molluscs, mites and insects. With the modernization of agricultural mechanization, the development of agricultural mechanization is complete, but due to the influence of geographical location and climatic conditions, it is subject to the trauma of drug

¹ Corresponding Author, Yuanbing LI, School of information engineering Sichuan agricultural university Sichuan, China; Email: 1046387092@qq.com.

resistance and heavy metal agents to a certain extent, which will have a certain impact on soil, surface water and other factors.

1.1. Research Status at Home and Abroad

With machine learning and deep learning technology, the convolution neural network established can make the whole process of input, output and integration easier. For example, Zhao and others use the partial least squares regression method to analyze the cucumber spectral map, introduce a variety of features such as entropy, contrast and homogeneity, and compare ANN (artificial neural network) and SVM (vector machine) for classification and recognition. The convolution neural network and cyclic neural network developed in recent years are also applied through the input layer, convolution layer, pooling layer, full connection layer and output layer, and some excellent frameworks Alex net and fast RCNN.[5] MLearned a new model through the migration of googlenet convolution neural network in 2016. Sladojevic et al. Developed a model based on caffe net on the Caffe framework, Based on image recognition, diseases and pests can be better judged and analyzed.

1.2. Subject Source and Main Research Status

1.2.1. Subject Source

This topic comes from the undergraduate scientific research interest project team, mainly studies the plant protection monitoring system, and is completed with the support of schools and teachers.

1.2.2. Research Status

Firstly, collect the data set, collect the pest pictures of corn and soybean, and extract and process the features of the images. Based on the pytorch dynamic framework, the data set is then subjected to image preprocessing operations such as labeling, wavelet algorithm and normalization. Then, the accuracy is obtained through the operation of feature and framework. Then, it is predicted according to the whole cycle of plant monitoring system, and then the model of the whole plant protection monitoring system is obtained. Computer image processing mainly includes the process of information capture, image preprocessing, image feature processing and obtaining results. Therefore, we can modify it in the whole stage of image processing, so as to improve the accuracy of image recognition and optimize the plant monitoring system.

2. Dataset Preparation

The traditional recognition pattern consists of five cases - data acquisition and analysis, data preprocessing, feature selection and sorting, convolutional neural network selection and optimization, result sorting, feedback and re optimization.

2.1. Building Data Sets

We constructed training sets and test sets of different kinds of samples, in which 180 images are the training set and 40 images are the verification set. Considering the deviation caused by the influence of data validation set on different effects, yolo3 model is used to process the data set.

The above pictures are the pictures of some training sets selected by us. The pictures are in JPG format and the resolution is 350×350, which meets the construction standards and training requirements of the data set. According to conventional standards, the ratio of training set to test set is about 9:1. At this time, 180 training sets and 20 test sets are selected, which roughly meet the general experimental standards. Most of the pictures come from the high-definition pictures of the network. In order to make the data results more reliable, split and mix the data set and test set of each type of picture, and then carry out training, which can improve the accuracy.

2.2. Data Preprocessing

In the actual shooting process, due to the objective and immeasurable factors such as light and humidity in the field, there are some distortions in the photos. In order to measure the color difference, the following RGB method is used.

As a widely used color standard in industry, RGB color mode mainly includes the changes of R (red), G (green) and B (blue) color channels and their different superposition degrees, so as to transform different colors. Through the selection, the target area can be continuously displayed.

The above figure 1 is a color cube commonly used in RGB color space. The origin represents black, and the farthest place from the origin is white. The gray value from black to white is distributed on the line from the origin to the farthest point, and different points in the cube represent other different colors. Therefore, in order to highlight the relevant target area, the superposition effects of green, red and blue are used for display.

Therefore, the gray histogram can be obtained by analyzing the RGB principle, and the conclusion can be obtained by using Python to analyze the gray histogram after pre-processing and post-processing.

Data preprocessing mainly uses image preprocessing based on wavelet algorithm. The basic idea of wavelet transform fusion algorithm: firstly, wavelet transform the source image, then merge the transform coefficients according to certain rules, and finally inverse wavelet inverse transform the wavelet transformed algorithm to obtain the fused image.[1]

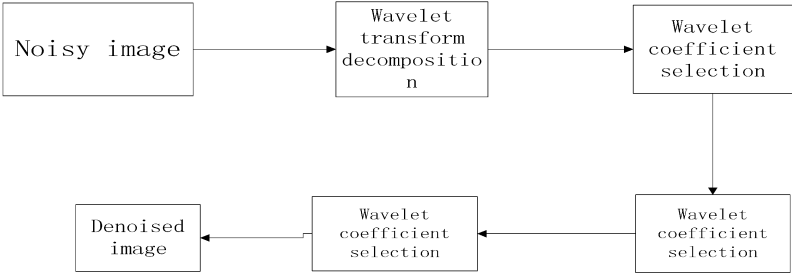


Figure 1. Wavelet change process

We carry out wavelet transform processing according to the images, and select the wavelet threshold, which is divided into hard threshold function and soft threshold function. In order to solve the one size fits all impact of the traditional hard threshold function, the soft threshold function will be less than 3σ . All wavelet functions are set to zero. After the preliminary processing of soft threshold, the image signal becomes smoother, and the adverse effect caused by local jitter is prevented. The choice of decomposition layers depends on the sampling frequency. We use three layers to transform the wavelet function and sym6 wavelet for operation. Decomposition can be performed in different dimensions. At this time, the threshold value is 3σ . If it is less than this threshold, it is the influence of noise. If it is greater than this signal, it is the influence of signal.

Then, after the local processing and the following processing of the photos we selected by Matlab tool, the following data can be obtained for recording. At this time, there are two kinds of recorded indicators, SNR and MSE signals, including noise signal $\alpha(1)$, including noise signal $\alpha(2)$, soft threshold function and new threshold function.[5]

Through the below table 1, we can analyze and get that the improved soft threshold function can eliminate the "burr". Although the data is not differentiable, after processing, we can get a new threshold function with a large decrease in MSE, which proves that the soft threshold function has a strong noise reduction ability for MSE, while the SNR signal has a weak noise reduction ability, and through the noise containing signal $\alpha(1)$ And including signal $\alpha(2)$ Finally, a new threshold function is obtained, which is obtained by their weighted average.[2-3]

2.3. Feature Selection and Preprocessing

After data preprocessing, we select the features for convolution neural network processing. At this time, the selected feature vector is very important. The feature selection is mainly based on the same feature on different images, and the labeling application is used for feature processing.

The main selection is the diseased part of corn and the included leaf part. For the diseased part, the more significant and special parts are mainly selected and the labels are noted to form multiple label block diagrams.

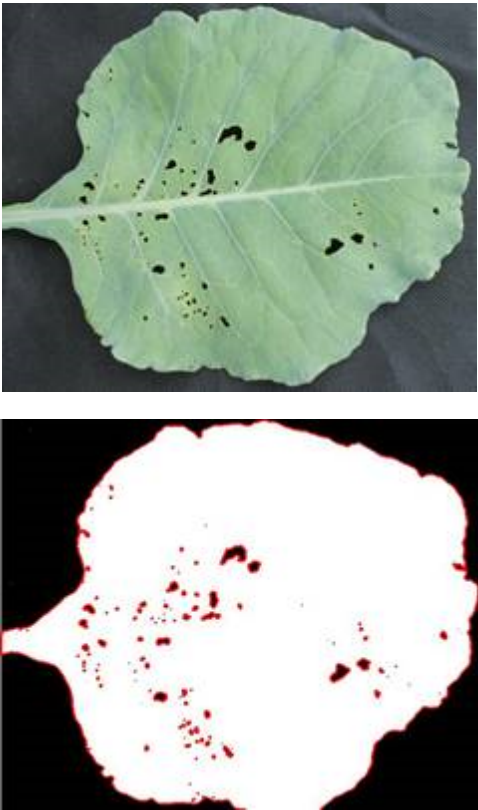


Figure 2. Leaf treatment

As shown in the figure 2 above, it is a processing method of label block diagram, in which the leaf parts other than the features are painted, so that the features can be clearly displayed, which is obviously helpful to the recognition rate.

By continuously selecting relevant label boxes in a picture, each label box contains the characteristics of associated diseases and pests, so as to continuously improve the corresponding accuracy.

As shown in the figure 3 below, it is a picture of feature selection in progress, in which each label box can be selected as many as two or more, and the XML file is generated through the size range of each label box[4]



Figure 3. Label box

We can observe the size and number of label frames of the files in the XML file generated by the JPG file, and then judge the quality of label frames in different images. The following lists the number of labels and corresponding size of label frames in a leaf of four different types to illustrate the characteristics of feature selection through the sampled data.

Table 1. Tag block diagram

Blade name and frame number	Label box size
Corn leaf spot 1 3	30^15
Maize leaf spot 1	20^17
Corn rust 1 2	10^42
Corn borer disease 1 2	40^21
Corn leaf spot 2	30^40
Maize small leaf spot 2	20^51
Corn rust 2 1	10^31
Corn borer disease 2 3	20^12

3. Convolutional Neural Network Training And Transfer Learning

3.1. Introduction to Convolutional Neural Networks

As one of the most widely used network models in deep learning, convolutional neural network (CNN) has attracted more and more attention and research. It is a special data network with similar network structure, which is called convolutional neural network. With the proposal of computer deep learning theory and the improvement and development of numerical computing equipment, convolutional neural network is mainly used in computer vision in the 21st century, such as pest recognition, target detection and recognition and natural language processing.

The convolutional neural network is shown in the figure 4 below, which is divided into convolution layer, pooling layer and full connection layer. The following structure is formed through different combinations (the figure 4 below is an example and will change according to the actual situation)

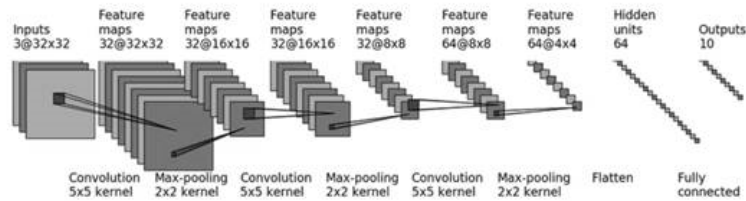


Figure 4. Convolutional neural network

The convolution layer uses multiple convolution kernels to extract the characteristics of the input data. Its three main parameters mainly include the size of convolution kernel, step size and filling function. The size of convolution kernel determines the complexity. The larger the convolution kernel is, the more complex the features are. Together, they determine the size of the output characteristic graph of the convolution layer, which is the super parameter of the convolution neural network. With the stacking of convolution layers, the size of the feature graph will gradually decrease. We can constantly change the size of the convolution kernel in the convolution layer to reach the corresponding feature graph degree.

The pooling layer is a layer that filters the convolution layer. The pooling layer contains the corresponding pooling function. Its function is to replace the result of a single point in the characteristic map with the characteristic map statistics of its adjacent areas. Controlled by the pooling size, step size and filling, LP takes the mean value in its pooling as the mean value pooling, and when the maximum value is taken as the maximum value pooling.[5]

The full connection layer is equal to the hidden layer of convolutional neural network and traditional feedforward neural network. The full connection layer is located in the last part of the hidden layer of the convolutional neural network. The feature map will lose the spatial topology in the full connection layer and be expanded into a vector. Through the excitation function, the extracted features are nonlinearly combined to obtain the output, which can be replaced by the global mean pooling.

3.2. Use the Relevant Yolo3 Neural Network

Yolov3 is an open source target detection project of Darknet, a niche deep learning framework. Darnet is short and fast, which can give full play to the function of multi-core processor and GPU parallel operation. Yolo3's fast detection is just suitable for video detection frames, and has high accuracy on small objects. Compared with other neural networks, Yolo has the following characteristics:

(1) The residual network is characterized by easy optimization and can improve the corresponding accuracy by increasing a considerable depth. The internal residual blocks use jump connection to alleviate the problem of gradient disappearance in the depth neural network, while the batch normalization and standard leakyreu parts are carried out behind dark net, It gives a nonnegative slope to all negative values.

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ \frac{x_i}{a_i} & \text{if } x_i < 0, \end{cases}$$

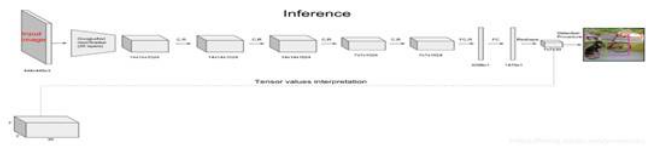


Figure 5. Algorithm change

(2) For backbone neural networks, Yolo has the following advantages (figure 5):

A. Feature extraction: compared with region proposal feature extraction in fast r-cnn (target detection) model, Yolo neural network can better distinguish the feature extraction mode of leaves and related leaf background, so it can better identify small and medium-sized leaf background.

B. Network prediction: yolov3 uses end-to-end prediction on the prediction picture, divides the whole picture into s * s regions, so that the blade will be divided into multiple modules, and makes a confidence judgment on the bounding box (prediction region) to determine whether the bounding box can be ignored. If it will not be ignored, it will be labeled.

C. The comprehensive and huge full connection layer uses a large number of convolution cores (more than 53). At the same time, the training speed on GPU is very fast, up to 15fps, fast accuracy and strong generalization ability.

	Type	Filters	Size	Output
1×	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
2×	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
8×	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
8×	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
4×	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 6. YOLOv3 core structure

As shown in the above figure 6, at the same time, there are three full convolution feature extraction layers on the 52, 26 and 13 dimensional models, corresponding to the convolutional set on the right, which is the internal convolution kernel structure of the corresponding feature extraction. The convolution kernel of 11 is used for dimensionality reduction, and the convolution kernel of 33 is used for the extraction of relevant features. All convolution layers are staggered connected. At the same time, the convolution layer

outputs a predict (prediction result), and the predicted result is regressed according to the confidence to obtain the final prediction result.

3.3. Blade Treatment Model

We put the blade related training set under the corresponding JPG file and XML file of Yolo model to start training. Because the number of feature processing and block diagrams under multiple files are different, we change the original range to 4 to be more compatible with the feature boxes.[6]

In the latest yolov3 model, the loss loss function is introduced.Initially, set the loss value loss of each image to 0, and the loss value loss of the total image to 0.Then the loss value loss can also be changed through the training value, so as to achieve a smaller loss value. Then check the GPU CUDA setting. If there is no GPU, run the relevant code with CPU.Clear the gradient, forward propagation, backward feedback, and then calculate the loss, where the calculated loss is equal to the loss of each picture. Carry out correlation superposition and summation, and divide by the number of all images to obtain the final loss value.

The next step is to carry out back-propagation and further stack the gradient. Finally, the total loss value can be obtained by stacking the loss of each item, and then carry out repeated training to reduce the loss rate. If the loss rate is smaller, it is proved that the training result is better.[7-8]

3.4. The Experimental Results are Shown and Discussed

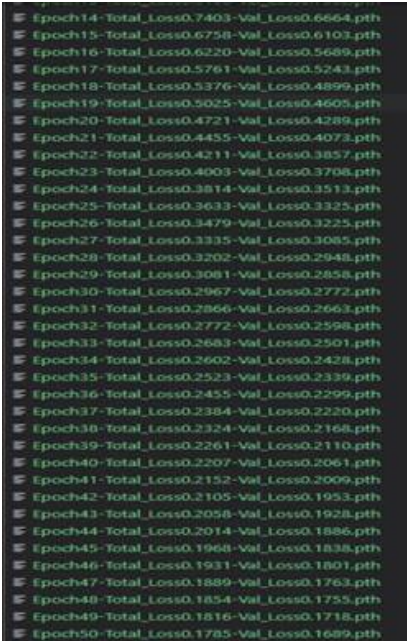


Figure 7. Loss iteration diagram

When calculating the parameter loss, it actually refers to the comparison between PRED and target. PRED refers to the prediction result and target refers to the target result.[8]

We can observe from the above figure 7 that after the training of yolov3 model, the loss degree will gradually decrease, so as to achieve the purpose of experimental training.

4. Conclusion

Through this experiment, we deeply realized the application of yolov3 neural network in agricultural pest identification. This neural network can continuously calculate the relevant loss value and relevant prediction value, so as to achieve the purpose of more accurate and rapid identification. Therefore, the improved yolov3 neural algorithm among the related algorithms can achieve a pioneering new chapter of convolutional neural network.

References

- [1] Kang Feilong et al 2020 Overview of image recognition application technology of multiple crop diseases and pests. *Jiangsu agricultural science* 48 (22): 22-27
- [2] Li Jiang et al 2021 Review on the application of machine learning in crop pest identification. *Henan agriculture*, 2021 (05): 49-50
- [3] Yu Hongtao et al 2019 Apple pest identification based on vgg-f dynamic learning model. *Science, technology and engineering* 19 (32): 249-253
- [4] Sun Jun et al 2017 Identification of various plant leaf diseases based on improved convolution neural network. *Journal of agricultural engineering* 33 (19): 209-215
- [5] Shen Kaiwen et al 2020 Design of visual pest monitoring system based on convolutional neural network. *Software guide* 19 (9): 122-126
- [6] Ji Xuewei et al 2020 Identification method of crop diseases and insect pests based on deep learning. *Nanfang agricultural machinery* 51 (23): 182-183
- [7] Tao Yi et al 2021 Design and application of intelligent agricultural pest identification system based on neural network. *Agriculture and technology* 41 (02): page 8-11
- [8] Song Yuqing et al 2020 Identification method of crop diseases and pests based on multi-layer EESP deep learning model. *Journal of agricultural machinery* 51 (08): 196-202